

**DEMARCATED DIGITAL CONTENT AND METHOD FOR CREATING
AND PROCESSING DEMARCATED DIGITAL WORKS**

RELATED APPLICATION DATA

[0001] This application is related to Applicants' patent applications entitled METHOD AND APPARATUS FOR TRANSFERRING USAGE RIGHTS AND DIGITAL WORK HAVING TRANSFERRABLE USAGE RIGHTS (Attorney Docket No. 111325-63), METHOD AND APPARATUS FOR ESTABLISHING USAGE RIGHTS FOR DIGITAL CONTENT TO BE CREATED IN THE FUTURE (Attorney Docket No. 111325-68), METHOD AND APPARATUS FOR DYNAMICALLY ASSIGNING USAGE RIGHTS TO DIGITAL WORKS (Attorney Docket No. 111325-66), METHOD AND APPARATUS FOR ASSIGNING CONDITIONAL OR CONSEQUENTIAL RIGHTS TO DOCUMENTS AND DOCUMENTS HAVING SUCH RIGHTS (111325-64), and METHOD AND APPARATUS FOR HIERARCHICAL ASSIGNMENT OF RIGHTS TO DOCUMENTS AND DOCUMENTS HAVING SUCH RIGHTS (111325- 65), which are being filed concurrently herewith, and are incorporated herein by reference in their entirety.

BACKGROUND

[0002] The subject invention relates generally to management of digital works and more specifically to demarcated digital content and a method for demarcating the content of digital works and processing the digital works.

[0003] One of the most important issues impeding the widespread distribution of digital works or documents (i.e. documents in forms readable by computers), via electronic means, and the Internet in particular, is the current lack of ability to enforce the intellectual property rights of content owners during the distribution and use of digital works. Efforts to resolve this problem have been termed "Intellectual Property Rights Management" ("IPRM"), "Digital Property Rights Management" ("DPRM"), "Intellectual Property Management" ("IPM"), "Rights Management" ("RM"), and "Electronic Copyright Management" ("ECM"), collectively referred to as "Digital rights

management (DRM)" herein. There are a number of issues to be considered in digital rights management: authentication, authorization, accounting, payment and financial clearing, rights specification, rights verification, rights enforcement, and document protection for example. U.S. patents 5,530,235, 5,634,012, 5,715,403, 5,638,443, and 5,629,980 disclose DRM concepts addressing these issues and are incorporated herein by reference.

[0004] In the world of printed documents, a work created by an author is usually provided to a publisher, which formats and prints numerous copies of the work. The copies are then sent by a distributor to bookstores or other retail outlets, from which the copies are purchased by end users. While the low quality of copying and the high cost of distributing printed material have served as deterrents to unauthorized copying of most printed documents, it is far too easy to copy, modify, and redistribute unprotected digital works. Accordingly, some method of protecting digital works is necessary to make it more difficult to copy them without authorization.

[0005] Unfortunately, it has been widely recognized that it is difficult to prevent, or even deter people from making unauthorized distributions of electronic works within current general-purpose computing and communications systems such as personal computers, workstations, and other devices connected over communications networks, such as local area networks (LANs), intranets, and the Internet. Many attempts to provide hardware-based solutions to prevent unauthorized copying have proven to be unsuccessful. The proliferation of high band-width "broadband" communications technologies will render it even more convenient to distribute large documents electronically, including video files such as full length motion pictures, and thus will remove any remaining deterrents to unauthorized distribution of digital works. Accordingly, DRM technologies are becoming a high priority.

00067754-053104
101E50-14529860

[0006] Two basic DRM schemes have been employed to attempt to solve the document protection problem: secure containers and trusted systems. A "secure container" (or simply an encrypted document) offers a way to keep document contents encrypted until a set of authorization conditions are met and some copyright terms are honored (e.g., payment for use). After the various conditions and terms are verified with the document provider, the document is released to the user in clear form. Commercial products such as IBM's CRYPTOLOPES™ and InterTrust's DIGIBOXES™ fall into this category. Clearly, the secure container approach provides a solution to protecting the document during delivery over insecure channels, but does not provide any mechanism to prevent legitimate users from obtaining the clear document and then using and redistributing it in violation of content owners' intellectual property.

[0007] Cryptographic mechanisms are typically used to encrypt (or "encipher") documents that are then distributed and stored publicly, and ultimately privately deciphered by authorized users. This provides a basic form of protection during document delivery from a document distributor to an intended user over a public network, as well as during document storage on an insecure medium.

[0008] In the "trusted system" approach, the entire system is responsible for preventing unauthorized use and distribution of the document. Building a trusted system usually entails introducing new hardware such as a secure processor, secure storage and secure rendering devices. This also requires that all software applications that run on trusted systems be certified to be trusted. While building tamper-proof trusted systems is a real challenge to existing technologies, current market trends suggest that open and untrusted systems such as PC's and workstations using browsers to access the Web, will be the dominant systems used to access digital works. In this sense, existing computing environments such as PC's and workstations equipped with popular operating systems (e.g., Windows™, Linux™, and UNIX) and

0986764.0311

rendering applications such as browsers are not trusted systems and cannot be made trusted without significantly altering their architectures. Of course, alteration of the architecture defeats a primary purpose of the Web, i.e. flexibility and compatibility.

[0009] U.S. patent 5,634,012, the disclosure of which is incorporated herein by reference, discloses a system for controlling the distribution of digital documents. Each rendering device has a repository associated therewith. A predetermined set of usage transaction steps define a protocol used by the repositories for carrying out usage rights associated with a document. Usage rights are encapsulated with the document content or otherwise associated with the document to travel with the document. The usage rights can permit various types of use such as, viewing only, use once, distribution, and the like. Rights can be granted based on payment or other conditions.

[0010] Fig. 1 is a block diagram of a conventional model for a system for the electronic distribution of digital works, which may include correspondence, books, magazines, journals, newspapers, other papers, software, audio and video clips, and other files objects, and the like. The phrase "digital work" as used herein refers to any type of element having content in computed readable form. "Content" as used herein refers to the viewable or otherwise usable portion of a digital work. Author 110 creates original content 112 and passes it to a distributor 120 for distribution. Ordinarily, author 110 is the creator of the content. However, the term "author" as used herein can be the creator, owner, editor, or other entity controlling the content or an agent (e.g. a publisher) of one of those entities. Also author 110 may distribute documents directly, without involving another party as distributor 120 and thus the author and distributor may be the same entity. However, the division of functions set forth in Fig. 1 is more efficient, as it allows author 110 to concentrate on content creation and not the administrative functions of distribution. Moreover, such a breakdown facilitates economies of scale by

09067754-053101

permitting distributor 120 to associate with a number of authors 110. Distributor 120 distributes digital works to user 130 upon request. In a typical electronic distribution model, the work is distributed as a document containing the content and associated usage rights in encrypted form. Distributor 120 encrypts the works with a random key and then encrypts the random key with a public key corresponding to user 130. Thus the encrypted work is customized solely for the particular user 130. User 130 is then able to use their private key to unencrypt the random key and use it to unencrypt and view the content at the work.

[0011] Payment for the work is passed from user 130 to distributor 120 by way of clearinghouse 150 which collects requests from user 130 and from other users who wish to sue a particular content. Clearinghouse 150 also collects payment information, such as debit transactions, credit card transactions, or other known electronic payment schemes, and forwards the collected payments as a payment batch to distributor 120. Of course, clearinghouse 150 may retain a share of the payment as a fee for the above-noted services. Distributor 120 may retain a portion of the batch payment from clearinghouse 150 for distribution services and forward a payment (for example royalties) to author 110. Distributor 120 may compile a bundle or batch of user requests for a single work before distributing the work. In such a case, a single instance of the encrypted work can be generated for unencryption by all of the requesting users 130.

[0012] Each time user 130 requests (or uses) content of a work, an accounting message is sent to audit server 140 which ensures that each request by user 130 matches with a document sent to user 130 by distributor 120. Accounting information is received by audit server 140 directly from distributor 120. Any inconsistencies are transmitted via a report to clearinghouse 150, which can then adjust the payment batches made to distributor 120 accordingly. This accounting scheme is operative to reduce the possibility of fraud in electronic distribution and to handle any time-

0986754-05101
T0750-154-9860

[0013] A first aspect of the invention is a digital work recorded on computer readable media comprising a content element including content data to be utilized by an end user, a content usage rights element associated with the content element and including data stored in addressable memory and specifying usage rights for the content element, and a flag element associated with the content element and having memory registers for saving demarcation flags.

BRIEF DESCRIPTION OF THE DRAWING

[0015] The invention will be described through a preferred embodiment and the attached drawing in which:

[0016] Fig. 1 is a block diagram of a conventional digital work distribution system;

[0017] Fig. 2 is a block diagram of the a storage device having a digital work in accordance with the preferred embodiment stored thereon; and

[0018] Fig. 3 is a schematic representation of the flag element of the preferred embodiment.

N182535.1

[0019] DRM techniques and systems permit digital works to be distributed electronically while allowing the content owner, such as the copyright holder, to control use of the content and receive remuneration for the content. However, there are still several factors that tend to prohibit widespread electronic distribution of many works, such as works having a large amount of content, valuable content, and the like. For example, many devices, such as pagers and personal digital assistants (PDAs), have limited memory and communications bandwidth. Therefore, in many cases it is not practical, or even possible to download a work in its entirety to such devices. Of course, the publisher could divide the intended work into plural documents of a smaller size. However, in such a case, the user would have to go through the purchasing and download procedure plural times to receive the entire intended work. Also, different devices have different memory sizes and communications bandwidth. Accordingly, the publisher would have to divide the document into documents of a size corresponding to the smallest memory and lowest bandwidth, i.e. the lowest common denominator, to ensure compatibility with all devices.

[0020] Consider a hypothetical situation in which a consumer orders a book or other digital work. If the digital work is too large to fit into the available memory of the user's device, then the download of the digital document will fail. Alternatively, if the work takes too long to download, the user may terminate the download. In either case, a sale may be lost, or the user may be charged for a work not received, depending on the fault tolerance of the distributor's system.

[0021] Applicant has discovered that portions of a work can be demarcated from the other portions thereof in a flexible and dynamic manner to be downloaded and used independently of other portions. Markers, i.e. flags, can be used to mark portions of the document that were downloaded and portions that were not downloaded. The flags and the content can be stored in the same memory device or in different memory device at different

096754.0341

[illegible][illegible][illegible]

more flags indicating which portions of content file 210 have been downloaded. If flag file 220 is updated whenever portions of content file 210 are downloaded, distributor 120 or clearing house 150 can monitor how much, and which portions, of content file 210 have been downloaded. This permits user 130 to be charged only for downloaded portions of content file 210 or to be charged only once for the entire content of content file 210 and permitted to download remaining portions of content file 210 at subsequent times. In such cases, flag usage rights 230 might only let distributor 120, clearing house 150, or another authorized party, manipulate flags 300 so that user 130 or others cannot tamper with the download record.

[0024] Fig. 3 illustrates the contents of flag file 220 of the preferred embodiment. Each box of Fig. 3 represents a memory register corresponding to an address in memory device 200, or a block of such addresses, corresponding to portions of content file 210. Flags 300 can be placed in the blocks, i.e. memory registers, to indicate the start or end, or any other aspect, of a demarcated portion of content file 210. In the preferred embodiment, flags 300 include start flags 310, and end flags 320 to respectively indicate the start and end of demarcated portions. In such a case, flags 300 can be represented by two integers to distinguish between corresponding start flags 310 and end flags 320. However, flags can be represented by a single integer if they are used merely as markers. For example, memory addresses of content file 210 between consecutive flags 300 can be demarcated. Alternatively, flags 300 can be any number of integers or in any form as required by the particular application and need to distinguish between flags 300. For example, as will become apparent below, flags 300 can serve many purposes and it may be desirable to distinguish between different types of flags. Also, a single flag can be used to mark content without a second corresponding flag.

[0025] In the preferred embodiment, the usage rights, in the form of content usage rights 240, are attached to content file 210 and are part of the

0986754-053101

[0027] Assuming the content in digital form, or a portion thereof, has been downloaded to a device, such as the user's personal computer, PDA, or the like, user 130 may begin reading the content, for example a book. However, user 130 most likely will not finish the book in one reading session. Accordingly, after the reading session, user 130 can manipulate flag file 220 to create flag 300 (serving as a digital bookmark) to mark the unread portion of the text. The manipulation can be accomplished by the user through the device standard interface using software running on the device or another coupled device or manipulation can be automatic. For example, when the user closes the application running the book or closes the file corresponding to the downloaded book content, flag file 220 can be manipulated accordingly. The newly created flag 300 is then saved in flag file 220 for the future use of user 130, or the use of other parties having authorization to access flag file

220 as specified by flag usage rights 230. User 130 can specify multiple flags for different locations in the book. The user can specify who will have authorization for access to flags 300, or the authorization can be supplied from a predetermined source, such as a database or directory.

[0028] User 130 can be more than one entity, in which case the multiple entities can have joint authorization for manipulation of flags 300 through flag usage rights 230. A multiple signature scheme or a voting scheme can be used to ensure proper authorization by joint users. Some users may be able exercise veto power. All such rights and rules can be implemented through flag usage rights 230 in a known manner. Access to flag usage rights 230, i.e. the ability to change the usage rights for flag file 220 can be granted in a similar manner.

[0029] Further, flags 300 can be in the form of calls, links, or other references to additional digital works 250 and thus can be used to permit the user to attach additional content data, such as one or more text documents, reminder notes, music, or multimedia pieces, to the content data in content file 210. Usage rights and flags can be assigned to a content file of the additional work 250 in a manner similar to content file 210. Accordingly, a hierarchical access scheme can be established. In other words, content is attached to other content via flags 300, which are attached to some other content via flags 300, in a pyramid or tree structure.

[0030] As, discussed above, one or more flags 300 can be used to demarcate a specific portion or portions of the content data in content file 210 to be selected, separated, cut, copied, printed, or otherwise processed in a desired manner. For example, in the case of the content of content file 210 being music, the frequency or volume of a portion of the music could be changed by demarcating that portion with flags 300. Further, the demarcated content portion can be filtered or compressed using an adaptive filtering or compression scheme. The demarcated portion also could be watermarked or

encrypted selectively. A demarcated portion of content corresponding to a textbook or a speech can be marked for translation to another language, using a translation engine or a human translator. Furthermore, the default currency used for payment for the content can automatically be changed in accordance with the translation language requested by the user.

[0031] Memory devices, specifically those in portable devices, are not always reliable. Therefore, a loss of data may occur, such as when the device operating system crashes, or when the device is subject to mechanical or electrical shock. In such a case, user 130 may desire to download the same content data or portions of content data to replace the content lost from memory. Of course, user 130 may have already paid for the content data. Flags 300 in flag file 220 can be used for aggregation and maintaining records of downloads and payments by user 130. For example, flag file 220 can be stored on the server of distributor 120 or clearinghouse 150 and flags 300 can demarcate portions downloaded and/or paid for by user 130. In such a case, flags 300 can be used for aggregating micropayments, i.e. plural relatively small payments, and charging the user at a future time, monthly or after downloading the entire content for example. Billing can be based on the credit limit user 130 applied for and obtained from distributor 120, clearinghouse 150, a credit card company, or other party financing the transaction. Also, flags 300 permit billing to be based on the portions of the content data purchased by user 130. Flags 300 can also be inserted automatically and be used to keep track of the reading habits (speed, frequency, and overall statistics) of user 130, based on the information gathered from many flags 300 during many sessions over a period of time.

[0032] In addition, flags 300 can be used to demarcate portions of a book or other content (by marking at least the beginning and the end of a portion of the content), from which a summary or abstract can be obtained automatically. Further usage rights and/or fees can be separately assigned to the abstract or summary as well. For example, flags 300 can demarcate

0906754-0340
T0F50-4529850

specific sentences or passages that convey the essence or a preview of the content for viewing as a summary. The demarcated portions can be merely parsed out of the content and presented as the summary. Alternatively, a neural network with fuzzy logic capability, or any other artificial intelligence engine, can be applied for proper organization and classification of the demarcated portions. Such an engine can be trained or tuned to organize the demarcated portions based on keywords or relevant topics. Examples of training algorithms for artificial intelligence are the iterative algorithm for obtaining a solution weight vector for linearly separable classes, and the training by back propagation method. In general, in a successful training session, the network error decreases with the number of iterations, and the procedure converges to a stable set of weights that exhibit only small fluctuations with additional training. Fuzzy logic engines (using membership functions and fuzzy set theory) provide more flexibility to the decision process. The decisions are not merely binary but can include fuzzy terms, such as "Low", "High", "Many", and "Approximately". This is useful for classification purposes. The obtained summary can be used for review purposes, for a database searched by a search engine, to direct the other users to the original book for further information. Demarcated words or other portions of the content data can be used for indexing a search engine using keywords/phrases selected by author 110 or publisher 120 as opposed to keywords determined by the search engine.

[0033] Furthermore, flags 300 can be used for the collection of educational course materials or assignments by a course instructor or other person setting curriculum. In particular, portions of the content can be demarcated, parsed out, and saved by distributor 120 (or given directly to students), for automatic collection and transmission of portions from different works, and for keeping track of charges to the students for those portions, according to the prices assigned for those portions. Flag file 220 can be customized for Teaching Assistants, students, or the like. For example, the Teaching Assistant's flag file 220 may include demarcated portions having quiz answers and teaching

aids. Of course such portions would not be demarcated by flag file 220 for students. It can be seen that one content file 210 can have plural flag files 220 or one flag file 220 can be associated with plural content files 210.

[0034] In addition, when traveling, the user can access a large amount of content stored on their home computer or elsewhere on their behalf using a hand-held device or a wireless device. Flags 300 can be used during unfinished downloads for large files to mark parts of the content for future downloads. The determination of required memory can be done automatically before a download procedure to streamline and optimize the downloading process by placing flags 300 in registers of flag file 220 demarcating content file 210 into portions of the desired size. Further, the system, such as a content server owned by distributor 120, can be configured to auto download the demarcated portions on a periodic basis.

[0035] Generally, flags 300 can be used for customization of content for each user by demarcating and keeping track of various portions of content. For example, each user can have one or more private and public libraries (a collection of books, book chapters, articles, or other works, selected by user 130), to which others can be granted access, based on the user's assignment of content usage rights 240. There may be more than one library, one for each subject (for example, one for a history collection and one for a scientific literature collection), in which case an automatic classification based on key words or neural networks can facilitate the selection of the appropriate default library. Personal libraries are folders for flag files 220 and can be distributed to others to mark and/or link content from various locations, such as various Web sties. The content of a library can be of any type or of any combination of types. The library allows the marked works to be easily correlated and retrieved for sale or other distribution.

[0036] User 130 can demarcate the name of books in a library (or alternatively, a small portion, a sample, or summary) by manipulating flags

09867754-053101
T07E50-45229860

300 and limiting content usage rights 240 by time, copies, or the like to encourage purchase of the content by peers for super distribution, i.e. redistribution by users 130. Group discounts referral points or monetary compensation can be given to the user for any redistribution.

[0037] Once the authorization granted by content usage rights 240 to use the content has expired, the content data can be automatically erased, or modified. For example, a the content data can be changed so a watermark or an intrusive message can appear on a rendered image, making the image useless or inconvenient. Or, for music, the frequency of demarcated ranges can be increased or silenced, making it unpleasant to listen to.

[0038] When content data is accessed from or stored on multiple servers or other devices, flags 300 can be used to keep track of servers for an optimized accessing scheme. For example, flags 300 can be used for identification and referral to a specific server for edge delivery of content over the Internet or any other network (as opposed to centralized content delivery), to solve the first-mile-bottleneck problem (related to traffic on the network and speed of delivery). Flag file 230 can be stored on or accessed by a central server which keeps track of flags 300 to allow various portions of content data to be demarcated for downloading from different servers based on load on each server or other characteristics and variables.

[0039] Another application of the preferred embodiment is in record keeping, such as laboratory notebooks or other notes kept by engineers, laboratory scientists and others. Conventionally, a laboratory scientist or engineer records test results, observations, and comments on a conventional paper lab notebook for future reference. The "lab notebook" is continuous, cannot easily be altered, is initialed and dated by the user, witnessed by others, and often is stored in a physically secure place. Such procedures render laboratory notebooks valuable as reliable evidence in legal proceedings, such as for establishing the date of recorded activity. Because

0986754 05101
T07E50 15429860

of the ability to tamper with digital records, digital records have not been utilized to a great extent as laboratory notebooks or other legal records.

[0040] To provide a comparable level of protection and integrity in a digital work, an electronic lab notebook can be provided, in which the text, figures, tables, or data can only be inserted at the end of the file. The text, figures, tables, or data can be changed, However, in such a case, all of the previous versions remain intact in the file for future inspection or auditing. The content data of the electronic lab notebook can be stored in content file 210 and previous versions can be demarcated by flags 300 to be hidden as a default for convenience and ease of review. The time and dates for each session can be recorded according to a secure and centralized clock, i.e. all entries are time-stamped, together with the electronic signatures, i.e. a code that can be attached to a document to uniquely identify an individual creating or modifying the document of the users and witnesses. Each session can be marked clearly by flags 300 and indexed in a database using an automatic search engine. The contents of the electronic laboratory notebook can be stored in content file 210 and the sessions, entries or previous versions are demarcated by flags 300 stored in flag file 220 and rights thereto are specified by content usage rights 240.

[0041] The overall work, or just content file 210, can be encrypted, or hash functions can be applied thereto, to insure the integrity of the data. A "hash value" is a number generated from a string of binary value. The hash is substantially smaller than the original data itself, and is generated by a formula in such a way that it is extremely unlikely that some other data will produce the same hash value. Hash values thus can be used to ensure that content has not been tampered with. The creator generates a hash value of the content, encrypts it, and stores the hash value with the content in content file 210. An authorized user can subsequently decrypt both the content and the hash value and produce another hash value from the content. The

0986754-053101
FOI 050-45249860

original and the new hash value can then be compared. If they're the same, there is a very high probability that the content has not been tampered with.

[0042] The right to inspect or audit the electronic laboratory notebook can be different from the right to use the same and can be effected through content usage rights 240. For example, the right to inspect may be granted only to a judge or an attorney. A copy of content file 210, content usage rights 240, flag file 220, and flag usage rights file 230 corresponding to the electronic lab notebook can be transmitted to a neutral, trusted third party for safe and redundant storage. The third party preferably will not have the right to inspect the content. Plural electronic lab notebooks can be linked together or merged, in which case the encryption or hash functions should be employed again to get the resultant or total of the contents secured as a whole. The order of the changes or mergers can be recorded and demarcated by flags 300 for future auditing purposes.

[0043] The preferred embodiment can be applied to all types of the digital works having any content, such as magazines, movies, multimedia, speech, software, and music. Further, the preferred embodiment can be used in any type of distribution model such as those based on subscriptions, club membership, pay-per-view, set-top-boxes, quantity of usage, usage period of time, expiration date, number of printouts, number of copies, purchase, view, try-before-you-buy, or rental.

[0044] As noted above, different parties may have different usage rights. Accordingly, the identity of a party can be ascertained through the use of biometrics (such as face recognition, iris recognition, eye recognition, fingerprint recognition, voice recognition, knuckle recognition, or hand recognition), signature analysis, Public Key Infrastructure (PKI) technology, password, device ID, smart cards, magnetic ID cards, and DRM schemes. Further, the usage rights can be granted only upon payment of a fee or other event. The content and flags of the preferred embodiment are separate files.

006754-0501
107550-45429860

However, the content and flags can be stored in any element and need not be separate files.

[0045] The distribution, accounting, and other functions of the distributor and clearinghouse can be accomplished by any party on any device. For example, the content can be rendered on an ebook reader or PDA in response to entry of a code or insertion of a smartcard into a reader and accounting can be accomplished when the digital work or accounting data is returned to a specific source. The division of tasks disclosed herein is only an example. Usage rights and or accounting data can be encapsulated with the digital work or can be stored separately. Code for rendering, decrypting, or otherwise permitting or limiting use of the content can be stored on any device or can be encapsulated with the digital work. Any distribution arrangement can be used with the invention and such arrangements can include any combination of devices, such as personal computers, servers, PDAs, and the like communicating with one another in any manner as is necessary to transfer the desired information.

[0046] The invention has been described through a preferred embodiment. However, various modifications can be made without departing from the scope of the invention as defined by the appended claims and legal equivalents.

0935754 053101
F0T50 4549960